Reviewer A Comments (Round 1)

For author and editor

Please see the attached PDF file for my comments.

This manuscript reports a comprehensive study of the seismicity in Cochabamba and the surrounding areas in Bolivia – from deployment of instruments, catalog building, to analyzing the seismicity. There is indeed a value in the work, however, I find that much improvement is needed on the overall structure of the manuscript as well as writing in general to make it read better and clear. There are also several technical details that need to be clarified and elaborated to avoid confusion. I hope the authors will find my comments listed below useful.

Comments on structuring the manuscript

- The introduction is well written, but I would merge the two paragraphs stating the objective of the study into one and make it clearer. Some of the details on the geology and tectonics can be moved to the discussion where you explain the seismicity to make the introduction more concise.
- A lot of the contents in the Results section should probably go into the Method section. There are also near duplicate statements in both the Method and the Results sections (especially on phase picking and phase association). An alternative approach is to make a section called "Event catalog building" and put everything from phase picking to magnitude estimation ordered by which comes first in the data processing workflow.
- I would also suggest having a look at papers published by other seismic network operators. For example, TexNet (https://doi.org/10.1785/0220180350), Oklahoma Geological Survey (https://doi.org/10.1785/0220190211), etc.

Comments on technical details

- **Magnitudes:** There are no clear explanations on how the magnitudes were calculated other than the mention of SEISAN in Lines 204 – 205. Were all magnitudes direct outputs from SEISAN? If that's the case, I wonder why there were no efforts on calibrating the local magnitudes for the region/network. Also, the frequency-magnitude plots in Figure 4b are somewhat strange because there is a single event with magnitude 4.4, absolutely no events between 3 and 4.3, and then suddenly the distribution follows the Gutenburg-Richter below magnitude 3. I am suspecting that the magnitudes other than that of the 4+ events were the direct results from SEISAN. If this is the case, I believe all the magnitudes are being under-estimated. A magnitude of completeness of 1 from a manually processed catalog is a surprisingly low value. If my understanding is correct, the authors should recalculate the magnitudes to make them consistent. If my understanding is incorrect, the authors should make the explanation clear to avoid any confusion.

- **The REAL associator:** This is a grid search method that involved many parameters such as the number of stations, phases, etc. The choices of parameters (and the reasons behind the choices, if possible) should be either mentioned in the manuscript or in the supporting information.
- **Phase picking:** It was unclear to me why the authors used an automatic workflow while the manual workflow already gave good results. Was this to complement the manual catalog? Or was this to correct analyst picks? This part of the manuscript currently reads as if the authors were testing the performance of this picker on their dataset, which doesn't necessary flow well with the theme.
- **Velocity model:** Much effort was put into developing a 1D velocity model but the motivation for doing this was unclear. The VELEST solution didn't change much from the Ryan et al., 2016 model. There is a mention of epicentral and depth errors, but it is wrong to refer to these errors as uncertainties. Were the different velocity models used to get an actual uncertainty estimate in event locations? If this was the case, the authors should clearly mention this.
- **Double-difference relocation:** This is usually considered as an important step, especially when trying to associate earthquakes with faults. I was wondering why the authors left this step out.

Minor comments

- Lines 140, 146, and 196: I would suggest using *workflow* instead of *protocol*.
- Lines 169 – 173: Threshold values of 0.5 and 0.55 are not necessarily high. I would rather say "following the convention".

Reviewer B Comments (Round 1)

For author and editor

This study used a local seismic network geometry with 11 stations to obtain a 6-month earthquake catalog created through both automatic and manual workflows. The manual catalog was processed by analysts using SEISAN, while the automatic catalog utilized PhaseNet for phase-picking, REAL for phase association and location using a 1D velocity model, and Hypo 71 to obtain the absolute location. Both manual and automatic catalogs were compared to produce a final catalog, with magnitudes computed using SEISAN. Additionally, this study shows a deep analysis of uncertainties considering the picks and catalog results, as well as the possible biases due to the use of a 1D velocity model. The study also presents the focal mechanisms of two significant earthquakes during the recorded period. Finally, it discusses the catalog and its relation to structures at depth.

I consider this research to be solid. However, I think there are some important items to consider regarding how the paper is approached in certain sections (please keep in mind the comments in the review attached). Additionally, there are general comments about some clarifications in specific lines, typos, figures, and references.

Review

1. **Catalog Methodology**: This study obtained a 6-month earthquake catalog created through both automatic and manual workflows. The manual catalog was processed by analysts using SEISAN, while the automatic catalog utilized PhaseNet for phase-picking, REAL for phase association and location using a 1D velocity model, and Hypo 71 to obtain the absolute location. Both manual and automatic catalogs were compared to produce a final catalog, with magnitudes computed using SEISAN.
    - I find the methodology for obtaining the catalog appropriate. However, the abstract does not effectively present the overall scope of the work described.
2. **Results**: The results are described in the following order:

3.1 Detection, Association, and Location:

- In this section, you mainly analyze detection and location performance. However, you do not address the association step, such as the number of picks your associator used to get locations or the number of picks lost during association. This analysis is crucial because the association step significantly impacts the automatic catalog. Even with good phase picking, a poor associator can negatively affect your catalog.

3.2 Local Uncertainties: **3.2.1 Pick Accuracy**:

- You already discussed pick accuracy in lines 189-196. It would be better to include these lines in this section to avoid redundancy, ensuring that the articulation of the figures and text remains coherent.
- Clarify what you mean by pick accuracy. In this section, you analyze the time differences between the manual and automatic catalogs. The term accuracy in machine learning typically refers to how often the model correctly predicts the outcome (correct predictions/all predictions).
- Regarding pick performance, lines 172-173 only mention checking the list of true positives and false positives (see my general comment on line 173). You could remove these lines from there and elaborate on your picks results here.

3.2.2 Velocity Model

3.2.3 Joint Relocation and Inversion of 1D Velocity Model

3.2.4 Testing the Implications of the Velocity Model

- Please correct me if I am wrong, but these last three subsections (3.2.2, 3.2.3, 3.2.4) aim to determine the confidence of the 1D velocity model in the study region, assessing its impact on source parameters (hypocentral location and origin time). If I understand correctly, I suggest combining these into a single subsection called "Velocity Model." In this subsection, describe the two approaches used for analysis: 1. The Vp/Vs ratio analysis for each station to review the model's lateral variation, showing possible location differences you can expect. 2. The results from VELEST to review potential biases using the velocity model and suggest epicentral uncertainties of around 2-3 km. In section 3.2.3, you discuss depth uncertainty; please refer to my comments on lines 260-280.

3.3 Focal Mechanisms

- Refer to my comments on lines 282-308.
- In your abstract, you state: “We also test the network's ability to resolve focal mechanisms of moderate to small events with a combined inversion of waveforms and polarities.” Be more specific here. In this section, you analyze only two earthquakes (ML 4.3 and ML 2.8). While this is good, it is not sufficient to generalize the network's ability to determine focal mechanisms for small and moderate events. Consider mentioning in the abstract that you reviewed the focal mechanisms of these two events to provide an overview of the mechanisms in the area.

3.4 Lateral Variations of Seismicity

- I am ok with your description in this subsection.

3.5 Seismicity and Its Relation with Structures at Depth

- I am ok with your description in this subsection.

General

1. **Line 39:** Adding a comma for clarity.
    - Change: “large uncertainties, which hinder their association with specific faults. We established a regional…”
2. **Line 41-42:**

    Here and over the complete paper: I think it’s important to highlight that you are using a pretrained deep learning model, unless you have retrained it for your zone.
3. **Line 45:** Minor change.

- Change: “hangingwall” to “hanging wall”

4. **Line 94-95:** There is an unexpected line break.
5. **Line 116:** Minor change.
    - Change: “of the PhaseNet” to “of PhaseNet”
6. **Line 130:**
Table S1 only, is not necessary to add “,supplemental material”
7. **Line 164:**
    - PhaseNet provides arrival times for both P and S waves, but it does not specify if it is Pg or Pn phases. So, I think you should clarify that, or even better, just mention it as P or S phases.
8. **Line 169-173:**
    - I consider you should provide more description about PhaseNet performance in terms if it effectively works well. If it missed some events, do you found an explanation? Maybe because a poor signal to noise ratio? . I know you mentioned the supplementary material, but once you used it you should describe better your results in this part
9. **Line 173:**
    - “The list of true positives and false positives is in the supplementary material S2a and S2b.” I think you are referring to Table S2. If that is correct, I do not fully understand your analysis of the catalog of picks. As you mentioned in the label, those are the picks after the association step. Therefore, you are not considering the complete catalog of picks, only those retained after the association step.
10. **Line 176-177:**
    - You are writing a general description of what an association algorithm does, but you are not mentioning how the REAL algorithm works and how it differs from other association algorithms. Furthermore, you could give a brief explanation of why you selected this association algorithm instead of other recent algorithms like GaMMA (Zhu et al.) or the association algorithm using GNNs (McBrearty, 2023).
11. **Line 205:**
    - You computed the local magnitude using SEISAN. However, it is ideal to provide a general understanding of how SEISAN computes local magnitude. Additionally, consider including information on the relevant physical hyperparameters. This information is useful because your catalog

could serve as a reference in your study area, making it important to detail this significant step, such as the magnitude estimation process.

12. **Line 214:**
    I think you are referring to the figure 5 c-d instead of figure 6 c-e.
13. **Line 218:**
    I think you are referring to the figure 5 f-h instead of figure 6 f-h.
14. **Line 216-217:**
    - I don't understand what you mean by 2x *0.25* VP = 3km. What is VP? Could you clarify that?
15. **Lines 240-244:**
    - These lines should be included in methodology and no in results.
16. **Line 243:** Minor change.
    - "… of 138 earthquakes. . Furthermore" to "… of 138 earthquakes. Furthermore"
17. **Line 243-244:**

    Furthermore, we opted to input the velocity model from Ryan et al. (2016) into VELEST.
18. **Line 244:**
    - "Table S4)" ? That table is not shown.
19. **Line 257:**
    - See section 3.2.3
20. **Line 265:**
    - Change "t0-depth trade off" to "time origin and depth trade-off"
21. **Line 260-280:**

    I understand you want to highlight that the depth of the earthquake depends on the epicentral distance to the closest stations and the accuracy of the velocity model. In cases of uncertainty in depth, the origin time will be affected, and vice versa. For example, in the Chapare Norte cluster, you aim to use the observed S-P delays at the closest station to estimate depth uncertainty, assuming correct epicentral distances and velocity model.

    However, I think this assumption is too general to be considered. It seems unnecessary to provide all the details to show a rough estimate of depth uncertainty. Wouldn't it be better to discuss the uncertainties provided by Hypo71 (ignoring the potential errors from the velocity model)? You have already demonstrated that the velocity model works in your area with Wadati plots and VELEST results

22. **Line 267-269:**
    - I am lost here because you mention the Chapar Norte cluster, and until this moment, I don't know where the cluster is because you have not drawn the seismicity that you got.
23. **Line 279-280:**
    - Additional analysis related to "….." and "…" have been performed, and the results …
24. **Line 282-308:**

    Keep in mind that you are describing the methodology here, not discussing your results. Therefore, you should include these lines in the methodology section, so you can focus directly on the results.
25. **Line 283:**
    - What does FMNEAR mean, and could you elaborate a little on this to provide an idea of what it does?
26. **Line 288-289:**
    - Do you mean you will call that earthquake the "Chaparo Norte Earthquake"?
27. **Line 289-290:**
    - Same comment as above but with the Cochabamba Centro earthquake?
28. **Lines 295-296:**
    - Additionally, the user can adjust filtering bands for each station and component individually and explore source depth with a user-defined depth step.
29. **Lines 325-326:**
    This is the largest earthquake recorded by the network in the period under review.

Data and code availability

When referring to web pages, I believe that the date of the last time access is confirmed should be mentioned.

References:

Line 479:

Center, E. (2014). Southern California earthquake center. *Caltech. Dataset.* – I do not know if the reference is correctly written, could you verify it?

Line 593-959

OSC, O. S. C. (2022). *20220518_Informe_sismico_04_2022_Provincia_Chapare_CB*. https://osc.org.bo/images/sismicos/20220518_Informe_sismico_04_2022_Provincia_Cha pare_CB.pdf – I do not know if the reference is correctly written, could you verify it?

Figures:

1. **Fig 1:** Could you specify the time period you are using in this catalog?
    - In the upper right figure, the depths of subduction are too small. I think you can show only the contours in intervals of 100 and increase the font size of the respective values.
2. **Fig 2.b:** You should add a text in your maps locating the Chapare region. Furthermore, in the label of your figure, you should mention that the regions Cochabamba, Punata, Sacaba, and Chapare are shown in both maps.
3. **Fig.** 3. Your global workflow does not show the use of HYPO701 to obtain the absolute location
4. **Fig.** 4. In the labels and captions, it's better to use Manual Catalog, Automatic Catalog. Keep in mind that you didn't only use PhaseNet to obtain your automatic catalog, REAL is also a key element in the results of your catalog.
5. **Fig 5**. "c. to e." It's better "c-e", "f. to .h" -> "f-h"
6. **Fig 6:**
    - A and C) If the background color already represents the S-P, you could change the color of the earthquakes to black to contrast with the background. At this moment, it is difficult to see the cluster because the color is very similar to the background.
    - A and C) What do you mean when you say "S-P time delay for 0 km error" - epicentral error?
    - B and D) They have the same description; they should be differentiable.
7. **Fig 7:** There are three figures in figure 7d, but they are not well-ordered because they overlap with figure e.
8. **Figures 7 and 8:** I suggest that figures 7d and 8d could be in the supplementary section. I also think you don't use them in your analysis.

9. **Figure 9:** The large and small blue stars correspond to the May 18th (ML 4.3 / Mw 4.1) and August 14th earthquakes (ML 2.5 / Mw 2.5), respectively.
10. **Figure 9:** What does Tbalanced mean?

Dear Andrea and Reviewers,

On behalf of the authors, I would like to thank you for the comments, requests and suggestions.

Do please find hereafter the answers to all comments in *Italic and Blue* text format.

Thank you for your time and constructive reviews that helped improve our article,

----
Reviewer A:
Please see the attached PDF file for my comments.
Recommendation: Resubmit for Review

This manuscript reports a comprehensive study of the seismicity in Cochabamba and the surrounding areas in Bolivia – from deployment of instruments, catalog building, to analyzing the seismicity. There is indeed a value in the work, however, I find that much improvement is needed on the overall structure of the manuscript as well as writing in general to make it read better and clear. There are also several technical details that need to be clarified and elaborated to avoid confusion. I hope the authors will find my comments listed below useful.
Comments on structuring the manuscript

- The introduction is well written, but **I would merge the two paragraphs stating the objective of the study into one** and make it clearer. Some of the details on the geology and tectonics can be moved to the discussion where you explain the seismicity to make the introduction more concise.

  Thank you. We merged the two paragraphs into one that sets out the objectives (line 113 to 125)

  "This study introduces the Proyecto Oroclino Boliviano Network (POBnet), which is the first earthquake monitoring system in the central region of the Bolivian orocline, encompassing the Cochabamba, Punata, Sacaba, and Chapare regions (Figure 2b). The network, established in 2022, comprises 11 seismic stations. Its principal objectives are to enhance the local earthquake detection capacity, improve the accuracy of earthquake location, determine new focal mechanisms, and associate earthquakes to the fault system surveyed by geologists. The study describes an "automated workflow" using a supervised Deep Neural Network (DNN) approach called PhaseNet (Zhu and Beroza, 2019) and a manual "analyst workflow" led by expert seismologists in order to estimate the capacity of the automated workflow. The study examines the effect of the velocity model on the earthquake locations and also explores the potential of moment tensor inversion for moderate to small earthquakes (Mw 2.5 and 4.1). Finally, it confronts the earthquake catalog with a geological balanced cross-section in order to draw seismotectonic implications."

- A lot of the contents in the Results section should probably go into the Method section. There are also near duplicate statements in both the Method and the Results sections (especially on phase picking and phase association). An alternative approach is to make a section called "Event catalog building" and put everything from phase picking to magnitude estimation ordered by which comes first in the data processing workflow.

We agree and follow the suggestion to improve the order within the section Methodology, we present a modified organization of section 2:

2. Methodology.

2.1. Network Deployment.

2.2 Workflow

2.3 Automatic Phase Picking and Association.

2.4 Velocity Model.

2.5. Focal Mechanism Using Near-Source Records (FMNEAR).

- I would also suggest having a look at papers published by other seismic network operators. For example, TexNet https://doi.org/10.1785/0220180350), Oklahoma Geological Survey (https://doi.org/10.1785/0220190211), etc.

Thanks for mentioning this reference that we read and now refer in the bibliography

Comments on technical details

Magnitudes: There are no clear explanations on how the magnitudes were calculated other than the mention of SEISAN in Lines 204 – 205. Were all magnitudes direct outputs from SEISAN? If that's the case, I wonder why there were no efforts on calibrating the local magnitudes for the region/network. Also, the frequency-magnitude plots in Figure 4b are somewhat strange because there is a single event with magnitude 4.4, absolutely no events between 3 and 4.3, and then suddenly the distribution follows the Gutenburg-Richter below magnitude 3. I am suspecting that the magnitudes other than that of the 4+ events were the direct results from SEISAN. If this is the case, I believe all the magnitudes are being underestimated.

A magnitude of completeness of 1 from a manually processed catalog is a surprisingly low value. If my understanding is correct, the authors should recalculate the magnitudes to make them consistent. If my understanding is incorrect, the authors should make the explanation clear to avoid any confusion.

Yes, the magnitude ML was calculated with SEISAN (and complemented with Mw determined by waveform inversion for the 2 largest earthquakes).
True, the magnitudes should be calibrated: the regional attenuation in Bolivia is probably different to the california attenuation used in Richter (1935 and modified with Hutton and Boore, 1987). Unfortunately, not many earthquakes during the experiment were big enough to be associated with a mb or a Mw. Given these

limitations and constraints, we differ the regional calibration of the magnitude seismicity to a future article.
Therefore, for the sake of continuity with the national observatory catalog (www.osc.org.bo) communicated to the public after the seismic alerts, we used the uncalibrated SEISAN
We now add information about the magnitude computation within the section Results and its text from 264 to 271 for the new version:

«The magnitude was computed using the original Richter (1935) formula for Southern California, with improvements from Hutton and Boore (1987) regarding the values of constants a, b, and c. To obtain the ML magnitude, a simulation of the Wood-Anderson seismometer is applied to the displacement traces from the horizontal components of a seismogram. These traces are then filtered using a 2 Hz high-pass 2-pole Butterworth filter. The equation used for the magnitude calculation is:

$$ML = Log10(Amp) + 1.11\ Log10(dist) + 0.00189*dist - 2.09 \quad (1)$$

Where Amp is the maximum amplitude in nm and dist (distance) in km., the value for "a" is 1.11, "b" is 0.00189 and "c" is -2.09.»

- **The REAL associator:** This is a grid search method that involved many parameters such as the number of stations, phases, etc. The choices of parameters (and the reasons behind the choices, if possible) should be either mentioned in the manuscript or in the supporting information.

  We now add complementary informations concerning the choices of parameters and their reasons (lines from 186 to 201):

  "Despite the availability of new seismic phase association algorithms, such as GENIE (Mc Brearty & Beroza, 2023) or GAMMA (Zhu et al., 2022), we have chosen to use the REAL algorithm (Zhang et al., 2019). In fact, this algorithm, which has been accessible to the public for over four years, has been subjected to rigorous testing in a multitude of studies (e.g. Ammirati et al. 2022, Derode et al. 2023), thereby substantiating its potential and reliability for the specific instrumental network in question. REAL combines the advantages of pick-based and waveform-based detection and location methods. Unlike waveform-based methods that use continuous seismic data to search for events in 3D space, REAL focuses on a smaller area around the station with the current initiating phase and seismic picks. We used a coarse grid (0.02° x 0.02° x 2 km) to generate travel time tables for each source station pair, based on the Ryan et al. (2016) velocity model (Table S3a).
  In order to ensure a good quality of the detected events based on the multiple picks of the PhaseNet detections, we imposed a minimum of 4 P and 2 S on 4 different stations to consider an event detection. Then 6 phases (with PhaseNet

probabilities higher than 0.5) on 4 different stations are needed to associate and pre-locate an event with REAL. We then use the Hypo71 routines (Lee and Lahr, 1972) to obtain the absolute earthquake location."

- **Phase picking:** It was unclear to me why the authors used an automatic workflow while the manual workflow already gave good results. Was this to complement the manual catalog? Or was this to correct analyst picks? This part of the manuscript currently reads as if the authors were testing the performance of this picker on their dataset, which doesn't necessary flow well with the theme.

  The POBnet experiment was first set up as part of the first author's PhD. The experiment is currently supported by the OSC and DASE. At the beginning of the experiment, the OSC could provide a senior analyst and a team of engineers and technicians in the field. After the first weeks, during which it covered the first significant earthquake cluster, the Cochabamba municipality and other state-related actors expressed their interest in establishing a long-term, possibly permanent network.

  We then sought funding to maintain the network of stations, which is still operational in the field, but we needed assistance in finding a budget to employ sufficient human resources. The primary objective of the work on the automatic catalog is to demonstrate that this experiment could be sustainable in the long term, even with a budget limited to instrument maintenance and telecommunications. The secondary goal is to ensure that we have a very homogeneous catalog of phase picks in the future, regardless of whether new analysts are involved or not.

- Velocity model: Much effort was put into developing a 1D velocity model but the motivation for doing this was unclear. The VELEST solution didn't change much from the Ryan et al., 2016 model. There is a mention of epicentral and depth errors, but it is wrong to refer to these errors as uncertainties. Were the different velocity models used to get an actual uncertainty estimate in event locations? If this was the case, the authors should clearly mention this.

  While presenting the early results of our experiments to a seismology workshop, we were asked to show that the 1D velocity model derived from Ryan et al. (2016) was suitable for the area. We first compared it to other models from similar regions, such as the frontal fold and thrust belt in Argentina. We then prepared Wadati plots to learn more about possible Vp/Vs variations. We also decided to perform a VELEST and Joint Hypocenter Determination analysis to obtain the seismic station time corrections and a 1D velocity model enhanced for the zone of interest. Initially, we designed an S-P sensitivity test to confirm our absolute locations and quantify any biases in the results that are confronted with the geological balanced cross sections.

  We now express this strategy in the text slightly modifying the wording in order to be *clearer (lines 289 to 337).*

- **Double-difference relocation:** This is usually considered as an important step, especially when trying to associate earthquakes with faults. I was wondering why the authors left this step out.

  This paper focuses only on the absolute location, looking at how regional seismicity is related with the main fault system. We know that relative relocation is useful for understanding which focal planes are activated during a seismic crisis. However, we will leave this for future studies as it is more complex and will require more detailed analysis,sensitivity tests and second order interpretations that will require more details

Minor comments

- Lines 140, 146, and 196: I would suggest using *workflow* instead of *protocol*.
  Done, «2.2. Workflow»
- Lines 169 – 173: Threshold values of 0.5 and 0.55 are not necessarily high. I would rather say "following the convention".

  We rewrite the sentences to be clear (lines 180 to 185)

Reviewer B:

This study used a local seismic network geometry with 11 stations to obtain a 6-month earthquake catalog created through both automatic and manual workflows. The manual catalog was processed by analysts using SEISAN, while the automatic catalog utilized PhaseNet for phase-picking, REAL for phase association and location using a 1D velocity model, and Hypo 71 to obtain the absolute location. Both manual and automatic catalogs were compared to produce a final catalog, with magnitudes computed using SEISAN. Additionally, this study shows a deep analysis of uncertainties considering the picks and catalog results, as well as the possible biases due to the use of a 1D velocity model. The study also presents the focal mechanisms of two significant earthquakes during the recorded period. Finally, it discusses the catalog and its relation to structures at depth.

I consider this research to be solid. However, I think there are some *important items to consider regarding how the paper is approached in certain sections* (please keep in mind the comments in the review attached). Additionally, there are general comments about some clarifications in specific lines, typos, figures, and references.
Recommendation: Revisions Required

Review

- **Catalog Methodology**: This study obtained a 6-month earthquake catalog created through both automatic and manual workflows. The manual catalog was processed by analysts using SEISAN, while the automatic catalog utilized PhaseNet for phase-picking, REAL for phase association and location using a 1D velocity model, and Hypo 71 to obtain the absolute location. Both manual and automatic catalogs were compared to produce a final catalog, with magnitudes computed using SEISAN. I find the methodology for obtaining the catalog appropriate. However, the abstract does not effectively present the overall scope of the work described.

  Abstract Edited:

  Located in the heart of the Bolivian orocline, the Cochabamba department and its two million inhabitants are exposed to frequent seismic activity. However, the tectonic structures causing these earthquakes remain poorly identified. Indeed, Bolivia's national seismological network does not optimally cover the area and the hypocentral locations of local earthquakes are therefore subject to large uncertainties, which hinder their association with specific faults. We established a regional network consisting of 11 broadband and short-period seismic stations, spaced approximately 20 km apart. This study highlights the initial 6-month seismic report, which involved an automated deep neural network-based seismic phase picking utilizing a pre-trained model. A thorough comparison with a manual catalog by seismic analysts is conducted for validation. Focal mechanisms of significant earthquakes are determined from full waveform inversion and polarities. Our preliminary results document midcrustal microseismicity located in the Main Thrust fault shear zone, and in its hanging wall, in a region affected by tectonic slivers and transverse faults impacting the sedimentary cover. These outcomes provide fresh insights into the fault system's seismogenic behavior and potential across the Bolivian orocline.

- **Results**:
- *3.1. Detection, association and location.*

The results are described in the following order: In this section, you mainly analyze detection and location performance. However, you do not address the association step, such as the number of picks your associator used to get locations or the number of picks lost during association. This analysis is crucial because the association step significantly impacts the automatic catalog. Even with good phase picking, a poor associator can negatively affect your catalog.

True. We modified this part , mentioning (1) the number of picks that was taken as an expected criteria and (2) the thorough review we did at each step (lines from 197 to 201 lines):

- *3.2 Local Uncertainties: 3.2.1 Pick Accuracy:*

You already discussed pick accuracy in lines 189-196. It would be better to include these lines in this section to avoid redundancy, ensuring that the articulation of the figures and text remains coherent.
Clarify what you mean by pick accuracy. In this section, you analyze the time differences between the manual and automatic catalogs. The term accuracy in machine learning typically refers to how often the model correctly predicts the outcome (correct predictions/all predictions).
Regarding pick performance, lines 172-173 only mention checking the list of true positives and false positives (see my general comment on line 173). You could remove these lines from there and elaborate on your picks results here.

We updated the paragraph on the pick accuracy. It is now only in section "3.2.1 Pick Precision" .

You are right, we are not talking about machine-learning here.We are discussing the precision of the resulting picks (minimum time between manual and automatic pics). To avoid confusion, we changed ¨Pick accuracy¨ by "Pick precision".

"In order to evaluate the precision of the picks obtained through our procedure, we conducted a comparison between the outcomes from manual and PhaseNet techniques across the entire dataset. This entailed the calculation of time differences between the analysts' and the automatic picks. The methods yielded comparable arrival times for P and S, with standard deviations of the differences lower than 0.26 seconds (0.20s and 0.26s for P and S waves, respectively, fig, 5c-e). Considering a theoretical maximum P velocity of 6 km/s, these statistical differences should not involve earthquake location differences higher than 1.5-2 km. This is corroborated by the histograms presented in Figure 6f-h, indicating that the majority of the hypocentral location differences between manual and automated methods are less than 2 km."

- *3.2.2 Velocity Model, 3.2.3 Joint Relocation and Inversion of 1D Velocity Model, 3.2.4 Testing the Implications of the Velocity Model*

Please correct me if I am wrong, but these last three subsections (3.2.2, 3.2.3, 3.2.4) aim to determine the confidence of the 1D velocity model in location and origin time). If I understand correctly, I suggest combining these into a single subsection called "Velocity Model." In this subsection, describe the two approaches used for analysis:

1. The Vp/Vs ratio analysis for each station to review the model's lateral variation, showing possible location differences you can expect.

2. The results from VELEST to review potential biases using the velocity model and suggest epicentral uncertainties of around 2-3 km. In section 3.2.3, you discuss depth uncertainty; please refer to my comments on lines 260-280. Refer to my comments on lines 282-308.

Thank you, we updated the text taking into account the recommendations. See hereafter (lines 202 to 222):

« 2.4 Velocity Model.

The 1D velocity model extracted for the Cochabamba region from Ryan et al. (2016) exhibits comparable Vp and Vs velocities to those documented in the literature for the eastern front of the Andes in northwestern Argentina (e.g., Ammirati et al., 2015; Venerdini et al., 2020; and Table S3b and S3c). The regional velocity model derived from the receiver function study by Ryan et al. (2016) suggests that the velocity model covering the Altiplano to Brazilian craton is more likely to be 2D in relation to the geological anisotropy found in the area. The use of a 1D model could result in the biasing of earthquake locations. This is to be expected at a large scale, but could also be sensitive at the smaller scale of our network.

Two approaches were considered in order to validate the velocity model proposed.

a. Vp/Vs derived from Wadati plot: The Wadati diagrams were determined for each seismic station (Figure S3). The Vp/Vs ratio estimated from this diagram is then compared with the Vp/Vs ratios determined by Ryan et al. (2016). We then assessed the potential impact of the local Vp/Vs on the seismic locations.

b. A joint relocation and inversion of the 1D velocity model was conducted. We used the joint hypocenter, velocity model, and station coefficient determination using VELEST (Kissling et al., 1994) and the Joint Hypocenter Determination method (Pujol, 1992). In order to conduct this analysis, a set of relatively well-constrained earthquakes was selected from the initial catalog, comprising approximately 439 mid-crustal events. These events were then tested against the velocity model proposed by Ryan et al. (2016) in order to evaluate any discrepancies and refine the parameters of our model.

As specified by the reviewer B the velocity model validations methods may go to Methods under the "2.4 Velocity Model" (small changes in the writing), this commentary also carries the following sub commentaries:

- **Line 260-280:** "I understand you want to highlight that the depth of the earthquake depends on the epicentral distance to the closest stations and the accuracy of the velocity model. In cases of uncertainty in depth, the origin time will be affected, and vice versa. For example, in the Chapare Norte cluster, you aim to use the observed S-P delays at the closest station to estimate depth uncertainty, assuming correct epicentral distances and velocity model. However, I think this assumption is too general to be considered. It seems unnecessary to provide all the details to show a rough estimate of depth uncertainty. Wouldn't it be better to discuss the uncertainties provided by Hypo71 (ignoring the potential errors from the velocity model)? You have already demonstrated that the velocity model works in your area with Wadati plots and VELEST results"

The location uncertainties provided by Hypo71 are not representative of the true location uncertainties. Indeed, Hypo71 determines the uncertainties by calculating the root mean square (RMS) of the residuals between the observed and theoretical travel times. Nevertheless, relying exclusively on RMS can be insufficient, especially in areas *with complex velocity structures, as it may not adequately represent the variability in seismic wave propagation*.

An additional step was thus introduced, based on the VELEST and Joint Hypocenter Determination (JHD) methodology, as described in Venerdini et al. (2020) and discussed by Derode et al. (2023).

The sensitivity analysis corroborates the conclusion that the uncertainties obtained in VELEST (±10% in velocity) ensure that the depth variations do not exceed 5 km.

Consequently, we propose that VELEST + JHD + Sensitivity Test gives an idea of the order of magnitude of the hypocentral location uncertainty and bias

- **Line 282-308:** Keep in mind that you are describing the methodology here, not discussing your results. Therefore, you should include these lines in the methodology section, so you can focus directly on the results.

We now include these lines in the methodology section and merge all implications of the velocity model in section 3.2.2

«3.2.2. Testing the implications of the velocity model.

In order to estimate this, we first plotted the Wadati diagrams for each seismic station (Figure S3). With the exception of SODC8 and 7, the VP/VS ratio measured for every station is close to the 1.75 considered in the initial velocity model. Given the values and average distance between events and stations in our network (~30 km), it can be reasonably assumed that potential velocity model differences would likely result in location differences of less than 3 km.

Moreover, to ascertain potential additional biases coming from the 1D velocity model, we conducted the joint hypocenter, velocity model and station coefficient determination of VELEST (Kissling et al., 1994), and the Joint Hypocenter Determination (Pujol, 1992). In order to achieve this, a set of relatively well-constrained earthquakes was selected from the initial catalog of events. This dataset was then filtered to include only earthquakes exhibiting an azimuthal gap of less than 180 degrees. Following this preselection, a subset of 138 earthquakes was obtained. The velocity model based on Ryan et al. (2016), Table S4, was then injected into VELEST as an initial model.

The best model derived from VELEST, falls within less than 10% of all models aforementioned (Table S3d). The P and S time corrections calculated for each station do not exceed +/- 1.5 seconds. The seismic station that presents the largest delays is SODC8, around +1.5 s while the other seismic stations show negative values smaller than -1.3 s. The distribution of the delays are consistent with the first order geological variations, SODC8 being on the eastern flank of the orogen, above shallow unconsolidated sedimentary units overthrusting the Chaco basin, while the stations to the west of the network lay on higher velocities and thicker tectonic units. (Figures S4a to c).

The epicenter localization uncertainties depend mostly on the seismic network coverage (primary and secondary azimuthal gaps, distance to the closest station, number of stations, Bondar et al. 2004), and are quite insensitive to velocity model errors (e.g. Bondar et al. 2004; Bondar & Storchak, 2011; Laporte et al. 2024). The VELEST inversion provides horizontal uncertainties around 2-3 km, and can be considered significantly lower than 5km.

However, depth estimation is typically less well constrained. The accuracy of depth estimation is contingent upon the epicentral distance to the closest stations and the accuracy of the velocity model (e.g. Gomberg et al. 1990; Bondar et al. 2004; Husen & Hardebeck 2010; Bondar & Storchak 2011; Laporte 2021; Letort 2014). Furthermore, the depth parameter may be subject to a trade-off with the estimation of the origin time during the localization procedure (e.g. Hussen & Hardebeck 2010; Letort 2014).

In order to overcome the issue of the time origin and depth trade-off, it is possible to rely on observed S-P delays on nearby stations (e.g. Derode et al. 2023; Gomberg et al. 1990; Koirala et al. 2023). For the "Chapare Norte" cluster (May 18, 2022), 62 pairs of P & S arrivals are picked on SODC7, the closest station, located at less than 5km from the cluster, all with very similar S-P delays (2 +/- 0.1 seconds). These S-P delays depend on the epicentral distances, on the velocity model, and on the focal depth. First assuming the epicenters and the velocity model are correct, these S-P delays can therefore provide an estimation of the earthquake depths. Derived from S-P delays, the average cluster depth is thus found at 13.2 +-1 km (Figure 6a - b), close to the average original catalog depth (12.5 km).

The VP/VS ratio used for this study is around 1.75, but locally, around SODC7, this value is lower, around 1.69, estimated from the Wadati diagram (Figure S3). On the basis of this local VP/VS ratio, with the same P-wave velocities, and assuming that the epicentral location is correct, the S-P delays are consistent with a deeper cluster at 14.9 +/- .1 km (Figure 6 c – d). Further analysis has been conducted, and the results are presented in the supplementary data material (Figure S5).»

- *3.3 Focal Mechanisms*
    - Refer to my comments on lines 282-308.

In your abstract, you state: "We also test the network's ability to resolve focal mechanisms of moderate to small events with a combined inversion of waveforms and polarities." Be more specific here. In this section, you analyze only two earthquakes (ML 4.3 and ML 2.8). While this is good, it is not sufficient to generalize the network's ability to determine focal mechanisms for small and moderate events. Consider mentioning in the abstract that you reviewed the focal mechanisms of these two events to provide an overview of the mechanisms in the area.

We slightly modify the abstract

Abstract edited:

Located in the heart of the Bolivian orocline, the Cochabamba department and its two million inhabitants are exposed to frequent seismic activity. However, the tectonic *structures causing these earthquakes remain poorly identified. Indeed, Bolivia's national seismological network does not optimally cover the area and the hypocentral locations of local earthquakes are therefore subject to large uncertainties, which hinder their association with specific faults. We established a regional network consisting of 11 broadband and short-period seismic stations, spaced approximately 20 km apart. This study highlights the initial 6-month seismic report, which involved an automated deep neural network-based seismic phase picking utilizing a pre-trained model. A thorough comparison with a manual catalog by seismic analysts is conducted for validation. Focal mechanisms of significant earthquakes are determined from full waveform inversion and polarities. Our preliminary results document midcrustal microseismicity located in the Main Thrust fault shear zone, and in its hanging wall, in a region affected by tectonic slivers and transverse faults impacting the sedimentary cover. These outcomes provide fresh insights into the fault system's seismogenic behavior and potential across the Bolivian orocline.*

- **3.4 Lateral Variations of Seismicity**

I am ok with your description in this subsection.

- **3.5 Seismicity and Its Relation with Structures at Depth**

I am ok with your description in this subsection.

- **General**
- **Line 39:** Adding a comma for clarity. Change: "large uncertainties, which hinder their association with specific faults. We established a regional…"

> *«...the hypocentral locations of local earthquakes are therefore associated to large uncertainties, which hinder their association with specific faults….»*

- **Line 41-42:**

Here and over the complete paper: I think it's important to highlight that you are using a pretrained deep learning model, unless you have retrained it for your zone.

We used a pre trained model called NCEDC, 2014 and now clearly mention it in the introduction and methodology section

- **Line 45:** Minor change. Change: "hangingwall" to "hanging wall"

Done

- **Line 116:** Change: "of the PhaseNet" to "of PhaseNet"

Done.

- **Line 94-95:** There is an unexpected line break.

Fixed.

- **Line 116:** Minor change.

Fixed

- **line 130:** Table S1 only, is not necessary to add ",supplemental material"

Fixed

- **Line 164:** PhaseNet provides arrival times for both P and S waves, but it does not specify if it is Pg or Pn phases. So, I think you should clarify that, or even better, just mention it as P or S phases.

We now clarify it by mentioning P and S waves.

- **Line 169-173:** I consider you should provide more description about PhaseNet performance in terms **if it effectively works well**. If it missed some events, do you found an explanation? Maybe because a poor signal to noise ratio? . I know you mentioned the supplementary material, but once you used it you should describe better your results in this part

- **Line 173:** "The list of true positives and false positives is in the supplementary material S2a and S2b." I think you are referring to Table S2. If that is correct, I do not fully understand your analysis of the catalog of picks. As you mentioned in the label, those are the picks after the association step. Therefore, you are not considering the complete catalog of picks, only those retained after the association step.

This answer is valid for the last two comments mentioned by the reviwer as LINE 169-173 and LINE 173, to answer the first one please refer to Line 171-176 of new document, and to answer the second one, please refer to Line 176 of new document. Fixed the typo error.

The present study has been designed with the aim of assessing the efficiency of the global automatic workflow, which encompasses Phasenet picking, REAL association, and relocation. Given the limitations in personnel at the OSC, the objective of this study was to develop and test a workflow that could achieve comparable results to those produced by analysts at the OSC. The utilization of our workflow, in conjunction with the parameterization of Phasenet and REAL, resulted in the generation of a catalog of events that exhibited a 40% increase in the number of events present when compared to the manual catalog. Additionally, all the events present within the manual catalog were successfully identified, thereby fulfilling one of the key criteria for the selection of our parameterization. That is the main reason why we are calculating the Precision, Recall and F1 score only on the phases used by REAL to detect and locate the events.

- **Line 176-177:** You are writing a general description of what an association algorithm does, but you are not mentioning how the REAL algorithm works and how it differs from other association algorithms. Furthermore, you could give a brief explanation of why you selected this association algorithm instead of other recent algorithms like GaMMA (Zhu et al.) or the association algorithm using GNNs (McBrearty, 2023).

  At the time of this study (the beginning of 2023), GAMMA and GNN's were only partially available and very new, with almost no studies testing their performances. We have chosen to use the REAL algorithm mainly because it has been accessible to the public for over four years and has been subjected to rigorous testing in many studies (e.g., Ammirati et al., 2022; Derode et al., 2023). Specially adapted to the large number of picks provided by Phasenet, REAL combines the advantages of pick-based and waveform-based detection and location methods, which is very useful in this type of study.

- **Line 205:** You computed the local magnitude using SEISAN. However, it is ideal to provide a general understanding of how SEISAN computes local magnitude. Additionally, consider including information on the relevant physical hyperparameters. This information is useful because your catalog this significant step, such as the magnitude estimation process.

  We added some information concerning the magnitude (Lines 264 to 272)

  «The magnitude was computed using the original Richter (1935) formula for Southern California, with improvements from Hutton and Boore (1987) regarding the values of a, b, and c. To obtain the ML magnitude, a simulation of the Wood-Anderson seismometer is applied to the displacement traces from the horizontal components of a seismogram. These traces are then filtered using a 2 Hz high-pass 2-pole Butterworth filter. The equation used for the magnitude calculation is:

  $M_L = Log_{10}(Amp) + 1.11\ Log_{10}(dist) + 0.00189{*}dist - 2.09$

  Where Amp is the maximum amplitude in nm and dist (distance) in km, the value for "a" is 1.11, "b" is 0.00189 and "c" is -2.09.»

- **Line 214:** I think you are referring to the figure 5 c-d instead of figure 6 c-e.
  Fixed
- **Line 218:** I think you are referring to the figure 5 f-h instead of figure 6 f-h.
  Fixed

- **Line 216-217:** I don't understand what you mean by 2x *0.25* VP = 3km. What is VP? Could you clarify that?, 2 x 0.25 x VP = 3 kilometers

  We changed the sentence.

  Line 254 to 263:
  «In order to evaluate the precision of the picks obtained through our procedure, we

conducted a comparison between the outcomes from manual and PhaseNet techniques across the entire dataset. This entailed the calculation of time differences between the analysts' and the automatic picks. The methods yielded comparable arrival times for P and S, with standard deviations of the differences lower than 0.26 seconds (0.20s and 0.26s for P and S waves, respectively, fig, 5c-e). Considering a theoretical maximum P velocity of 6 km/s, these statistical differences should not involve earthquake location differences higher than 1.5-2 km. This is corroborated by the histograms presented in Figure 6f-h, indicating that the majority of the hypocentral location differences between manual and automated methods are less than 2 km.»

- **Lines 240-244:** These lines should be included in methodology and no in results.

  We have decided to maintain these lines within section 3.2.2, because it relates with the quantity of earthquakes taken into account for the VELEST algorithm, it is worth keeping in this section for the global understanding of results. (Lines 289 to 337)

- **Line 243:** Minor change. "… of 138 earthquakes. . Furthermore" to "… of 138 earthquakes. Furthermore"

  Fixed

- **Line 243-244:** Furthermore, we opted to input the velocity model from Ryan et al. (2016) into VELEST.

  FIxed

- **Line 244:** "Table S4)" ? That table is not shown.

Done

- **Line 257:** See section 3.2.3

Done

- **Line 265:** Change "t0-depth trade off" to "time origin and depth trade-off"

Done

- **Line 267-269:** I am lost here because you mention the Chapar Norte cluster, and until this moment, I don't know where the cluster is because you have not drawn the seismicity that you got.

We mention Chapare in the Figure.

- **Line 279-280:** Additional analysis related to "….." and "…" have been performed, and the results …

Fixed

- **Line 282-308:** Keep in mind that you are describing the methodology here, not discussing your results. Therefore, you should include these lines in the methodology section, so you can focus directly on the results.

  We now include these lines in the methodology section and merge all implications of *the velocity model in section 3.2.2*

- **Line 283:** What does FMNEAR mean, and could you elaborate a little on this to provide an idea of what it does?

We add a description of FMNEAR from lines 223 - 236

- **Line 288-289:** Do you mean you will call that earthquake the "Chaparo Norte Earthquake"?

There are no towns near the epicenter, we therefore named the earthquake from the name of the region in which it falls

- **Line 289-290:** Same comment as above but with the Cochabamba Centro earthquake?

Yes, because it was at the central part of the basin

- **Lines 295-296:** Additionally, the user can adjust filtering bands for each station and component individually and explore source depth with a user-defined depth step.

Done

- **Lines 325-326:** This is the largest earthquake recorded by the network in the period under review.

Done

- **Data and code availability**

When referring to web pages, I believe that the date of the last time access is confirmed should be mentioned.

- **References:**
- **Line 479:**

Center, E. (2014). Southern California earthquake center. *Caltech. Dataset.* – I do not know if the reference is correctly written, could you verify it?

We checked and replace the citation with
NCEDC, 2014. Northern California Earthquake Data Center. UC Berkeley Seismological Laboratory. Dataset.

- **Line 593-959**

We re do the citation according to «https://columbiacollege-ca.libguides.com/apa/websites»

OSC. (2022, May 18). Informe_sismico_04_2022_Provincia_Chapare_CB. https://osc.org.bo/images/sismicos/20220518_Informe_sismico_04_2022_Provincia_Chapare_CB.pdf

- **Figures:**
- **Fig 1:** Could you specify the time period you are using in this catalog? In the upper right figure, the depths of subduction are too small. I think you can show only the contours in intervals of 100 and increase the font size of the respective values. →

We optimized the slab inset figure.
We think we have a charged map to add more text, we write the catalog period in the labeling.

- **Fig 2.b:** You should add a text in your maps locating the Chapare region. Furthermore, in the label of your figure, you should mention that the regions Cochabamba, Punata, Sacaba, and Chapare are shown in both maps.

Done

- **Fig. 3**. Your global workflow does not show the use of HYPO701 to obtain the absolute location

Done

- **Fig. 4.** In the labels and captions, it's better to use Manual Catalog, Automatic Catalog. Keep in mind that you didn't only use PhaseNet to obtain your automatic catalog, REAL is also a key element in the results of your catalog.

Done

- **Fig 5**. "c. to e." It's better "c-e", "f. to .h" -> "f-h"

Done

- **Fig 6:** A and C) If the background color already represents the S-P, you could change the color of the earthquakes to black to contrast with the background. At this moment, it is difficult to see the cluster because the color is very similar to the background.

Done

The dots are "observed S-P delays "

A and C) What do you mean when you say "S-P time delay for 0 km error" - epicentral error?

We re write the Figure 6 labeling, however, we explain what we ment:

By '0-km error', we actually meant that Figure 6a shows the S-P delays according to their epicentral distance and their catalog depth, assuming that the epicentral distances found when locating the earthquakes are correct (without any location error, that would made us overstimate or underestimate the estimated epicentral distances).

B and D) They have the same description; they should be differentiable.

Done

- **Fig 7:** There are three figures in figure 7d, but they are not well-ordered because they overlap with figure e.

Done

- **Figures 7 and 8:** I suggest that figures 7d and 8d could be in the supplementary section. I also think you don't use them in your analysis.

We prefer maintaining these figures in the main text because the quality of a waveform inversion can hardly be appraised without showing the modeling of the actual waveforms.

- **Figure 9:** The large and small blue stars correspond to the May 18th (ML 4.3 / Mw 4.1) and August 14th earthquakes (ML 2.5 / Mw 2.5), respectively.

Done

- **Figure 9:** What does Tbalanced mean?

Sorry for the typo.
We selected the "Balanced cross section" of Mc Quarrie.

Reviewer B Comments (Round 2)

The authors have made substantial revisions to the order of the paper, addressing the most significant suggested changes. Although there are a few remaining comments (see them below), overall, I am satisfied with the scientific content presented in this research.

- **Line 85:** Remove white space.
- **Line 154:** Savvaidis
- **Line 163:** Associated and Located
- **Line 177:** Remove whitespace
- **Lines 265 – 266:** In a few words, explain the physical meaning of a, b, and c. Otherwise, the reader might be confused.
- **Line 266-267:** "To obtain the ML magnitude, a Wood-Anderson filter was applied to the multicomponent signal."
- **Line 290:** "In order to estimate the implications of the velocity model"

Reviewer C Comments (Round 2)

For author and editor

I'm reviewing the revised version of this manuscript directly. Previous reviewers have raised comments about the writing / structure of this paper. I agree with them, and the current version still has much room for improvement. For example, in the abstract, the key point of this study seems to be understanding the fault structures related to the seismic hazard, whereas in the later part of abstract, no related results are stated. The only finding seems to be the existence of mid-crust seismicity, but the authors describe it as "our preliminary results": if you can further improve the results or can further validate it, why not finish this work before submitting? Moreover, also from the abstract, I don't see much interpretation on the observations. You really need to tell us what does it mean to observe this seismicity pattern. Besides the issues on presentation, I also have some comments on technical part:

1. From Fig 4, it is obvious that the manual and automatic catalog is inconsistent for M>1. If you are using a merged catalog of these two, also add the FMD for the merged catalog, which is used in the discussions.
2. Fig 4, the Mc for manual catalog should be around 1.3, adopting Max curvature criteria.
3. The color for dots in Fig 4a is not explained in the caption. If it denotes the magnitude, no need to use different colors.
4. Fig 8, the waveform fitting is not very good. Any idea on that? Please add analysis on quality control in the main text. Also, for such small events, using polarity-based algorithms, e.g. HASH, is usually more stable. It is worth doing such validation, because I can tell from Fig 9 that the strike of this nodal plane is not consistent with local block movements.
5. In your location, do you consider the station elevation? From Fig 9, I think the elevation is not ignorable.
6. The title in the main text and supplemental material are different

Dear Andrea and Reviwers

Please find enclosed, with that letter, the revised version of our manuscript «Unveiling midcrustal seismic activity at the front of the bolivian altiplano, Cochabamba region» by Fernandez G.A. et al.

In line with the recommendations in your last email, we addressed the points raised by referees (in red hereafter), answering their remarks point by point in *Italic and light blue* color and slightly modifying the figure 4b, following the remarks of reviewer C.

We hope that this revised manuscript follows the guidelines of Seismica, and will be acceptable, both in substance and form. We stand ready to accommodate eventual changes following your advices/recommendations on these points as well as any additional requirements within short delays.

Sincerely,

Gonzalo Fernandez

Reviewer B:

- **Line 85:** Remove white space.

  Resolved.

- **Line 154:** Savvaidis

  Resolved, modification done.

- **Line 163:** Associated and Located

  Resolved, modification done.

- **Line 177:** Remove whitespace

  Resolved, modification done.

- **Lines 265 – 266**: In a few words explain the physical meaning of a, b and c.

  We now mention « values of the three constants associated to the formula, i.e. associated to the geometric spreading, the attenuation and a correction.»

- **Line 266-267:** "To obtain the ML magnitude, a Wood-Anderson filter was applied to the multicomponent signal."

  Resolved, modification done.

- **Line 290:** "In order to estimate the implications of the velocity model"

  Resolved, modification done.

Reviewer C:

I’m reviewing the revised version of this manuscript directly. Previous reviewers have raised comments about the writing / structure of this paper. I agree with them, and the current version still has much room for improvement. For example, in the abstract, the key point of this study seems to be understanding **the fault structures related to the seismic hazard, whereas in the later part of abstract, no related results are stated.**

The only finding seems to be the existence of mid-crust seismicity, but the authors describe it as “our preliminary results”: if you can further improve the results or can further validate it, why not finish this work before submitting?

Moreover, also from the abstract, **I don’t see much interpretation on the observations.**

You really need to tell us what does it mean to observe this seismicity pattern. Besides the issues on presentation, I also have some comments on technical part:

We recognize that the implications of the work in term of earthquake hazard were not detailed within the abstract.

We now improve the visibility of the conclusions of the implications mentionning in the abstract that :

«The seismic activity appears to be concentrated in the Main Thrust fault shear zone, and in its hangingwall, a stack of tectonic slivers affected by transverse faults. The depth of the events beneath the network is shallower than was previously thought. Seismic clusters in the Main Thrust System below the toe of the high topography are likely caused by strain and stress build-up on the frontal decollement. These results therefore provide new insights and hypothesis into the seismogenic behavior and potential of the fault system across the Bolivian orocline.»

In addition, we added a few references to previous studies in the discussion and comparisons with the results published in previous studies

«We interpret that part of this seismicity is a consequence of the activity of the main thrust system, as previously suspected in southern Bolivia and in Argentina (e.g. Isacks, 1988; Brooks et al., 2011; Weiss et al., 2016; McFarland et al., 2017; Figueroa et al., 2020; Ammirati et al., 2022).»

«Finally, evidence of reverse faulting activity is observed at mid-crustal depths, approximately 50 km from the surface trace of the most frontal thrusts beneath the toe of the

high topography in the Chapare region (Figure 9). We interpret this cluster as indicative of the down-dip extension of a 50 km-large partially or fully locked fault segments of the main active thrust system. This distance is less than that of the fully locked fault zone proposed further south along the southern branch of the orocline on the basis of the GPS velocity field. (Brooks et al., 2011; Weiss et al., 2016). In our region of interest, the Chapare cluster may be the result of persistent stress build-up at the downdip-end of a shorter locked fault zone, specific to the core of the orocline.»

- From Fig 4, it is obvious that the manual and automatic catalog is inconsistent for M>1. If you are using a merged catalog of these two, also add the FMD for the merged catalog, which is used in the discussions.

We do not use a merged catalogue, the intention of the graphics was (1) to compare first the automatic catalogue with the manual catalogue (2) to check the inconsistencies (3) to work with the automatic catalogue after manual validation.

In order to be specific we change the figure 4b to present the Frequency Magnitude Distribution (FMD) and the completeness magnitude (Mc) computation with Maximun Curvature Method (MAXC) which dentifies Mc as the point of maximum curvature on the FMD, making it ideal for initial assessments of seismic catalogs.

- Fig 4, the Mc for manual catalog should be around 1.3, adopting Max curvature criteria.

We now use a maximum curvature approach to determine the Mc. We obtain Mc=0.8. We report it on Figure 4b.

- The color for dots in Fig 4a is not explained in the caption. If it denotes the magnitude, no need to use different colors.

Yes, the color information denotes the magntiude, which is already expressed on the Y-axis and by the size of the circle. This information is therefore redundant. We modified this plot.

- Fig 8, the waveform fitting is not very good. Any idea on that? Please add analysis on quality control in the main text. Also, for such small events, using polarity-based algorithms, e.g. HASH, is usually more stable. It is worth doing such validation, because I can tell from Fig 9 that the strike of this nodal plane is not consistent with local block movements.

We recognize that waveform fitting in Figure 8 was rather poor for some stations. We explain this by the fact that, given the small size of the earthquake, waveforms have to be modeled at *relatively high frequencies (0.35 to 1 Hz). In this frequency range, as distance increases, the waveforms may reflect more complex wave propagation not accounted for by a simple 1D velocity model.*

In addition, the signal-to-noise ratio deteriorates as distance increases. As could be seen from Figure 8d of the manuscript, the worst fits are for the most distant stations (SOD2, SOD6, SOD7, SODX, and to some extent SOP0), while the fit is best for the four closest stations (SOP5, SOP4, SOP3, SOP1). We had chosen to keep and show the modelling of the most distant stations, with the aim of testing the entire network (local temporal and regional permanent stations) for small earthquakes.

Taking into consideration the remark of the reviewer, we could obtain a solution displaying a better overall waveform fitting by using only the four closest stations in the waveform inversion part. ***Figure 8 new*** *below is the new Figure 8 of the revised manuscript. As can be seen, the result in terms of focal mechanism and depth is almost the same.*

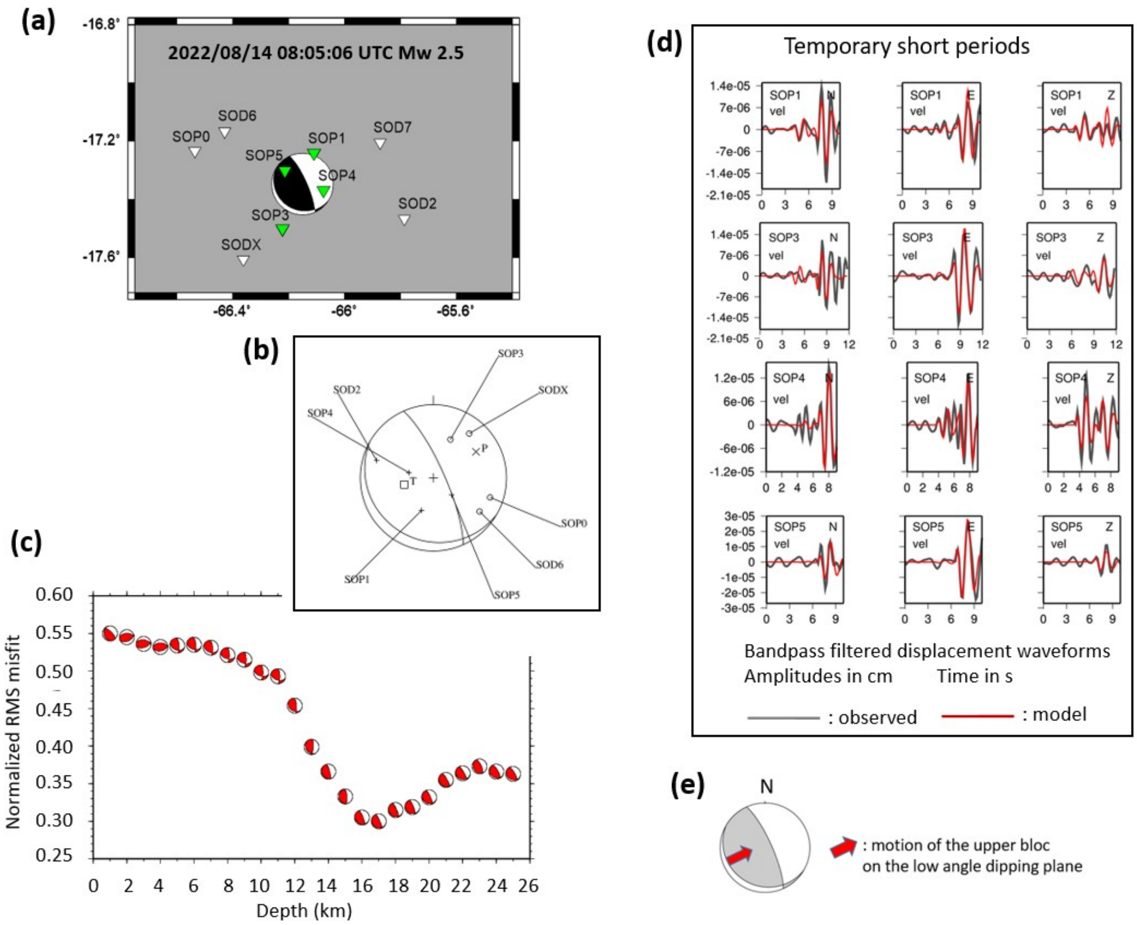

Figure 8 new. *Result of the joint inversion of first motion and waveform data for the 2022/08/14 08:05:06 UTC Mw 2.5 event, keeping only the four closest stations for the waveform part.*

- Also, for such small events, using polarity-based algorithms, e.g. HASH, is usually more stable.It is worth doing such validation

  *Given the small number of first motion data for this small event (at larger distances the low signal/noise ratio prevents from reading the polarity), we do not pretend that the polarities by themselves constrain the FM solution. As shown by **Figure S7**, many different FM solutions can explain all the polarities.*

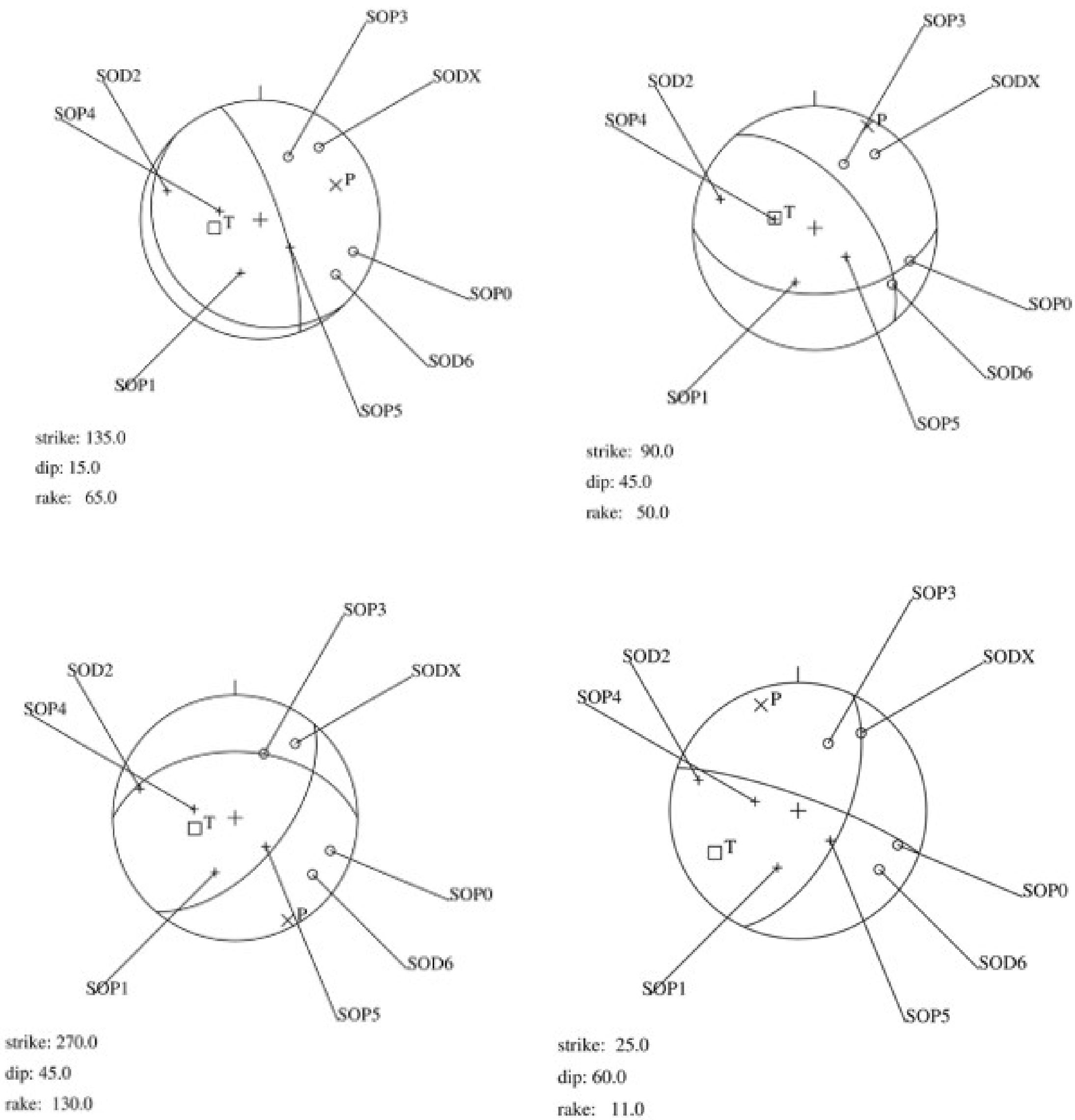

Figure S7: four different focal mechanisms solutions explaining equally well the first motion data, for the 2022/08/14 08:05:06 UTC Mw 2.5 event.

Certainly, if we did not have the waveforms, an alternative method like HASH would be quite appropriate. However, the combination of polarities and waveforms in a joint inversion greatly improves the stability of the solution, as we show below.

Our approach (FMNEAR, Delouis 2014) is also based on a grid search, combined with simulated annealing, enabling an exploration of the space of possible solutions for the focal mechanisms (FM).

*In **Figure S8** the best strike, dip and rake solutions from the inversion of polarities alone are plotted as a function of their RMS misfit value. As can be seen, a wide range of strike, dip and rake values is possible, explaining all polarities (RMS = 0), except for negative rakes corresponding to normal faulting solutions. This confirms that the focal mechanism is not constrained by the polarities alone.*

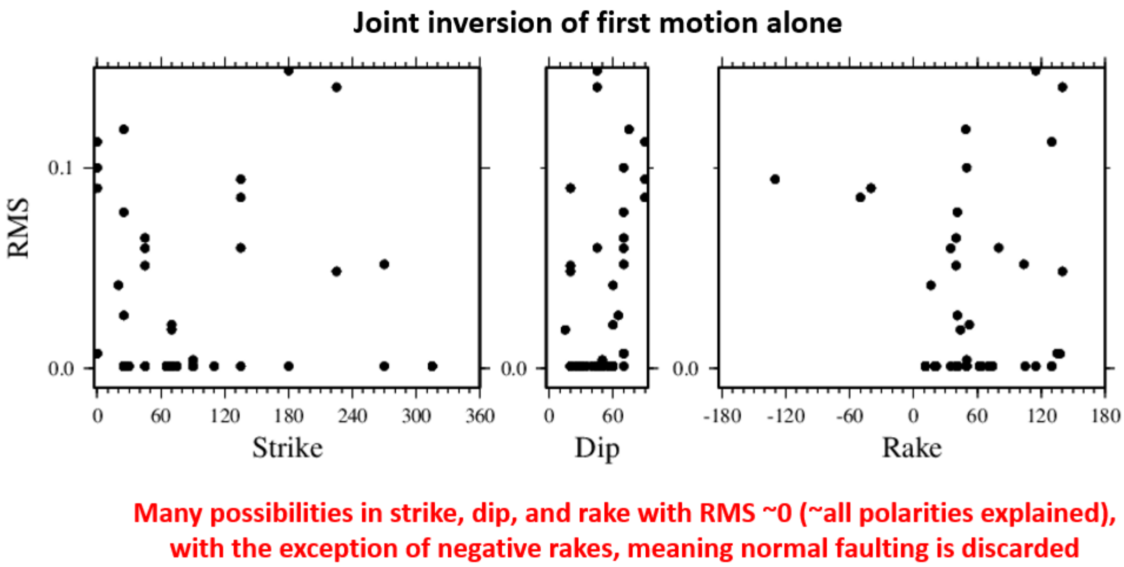

***Figure S8.** Best solutions found at the end of the various grid search and the simulated annealing combinations for the inversion of first motion data alone, for the 2022/08/14 08:05:06 UTC Mw 2.5 event.*

*When we invert the first motion and waveform data jointly, the solution becomes much better constrained, as shown by **Figure S9**, which is similar to **Figure S8** except that it is obtained by the joint inversion. The RMS values are higher because the waveforms are never perfectly matched due in particular to the inadequacy of the* *velocity model and the presence of residual noise in the data despite the filtering. This time we obtain two groups of solutions, named A and A', which in fact correspond to the same mechanism. The solution is therefore well constrained.*

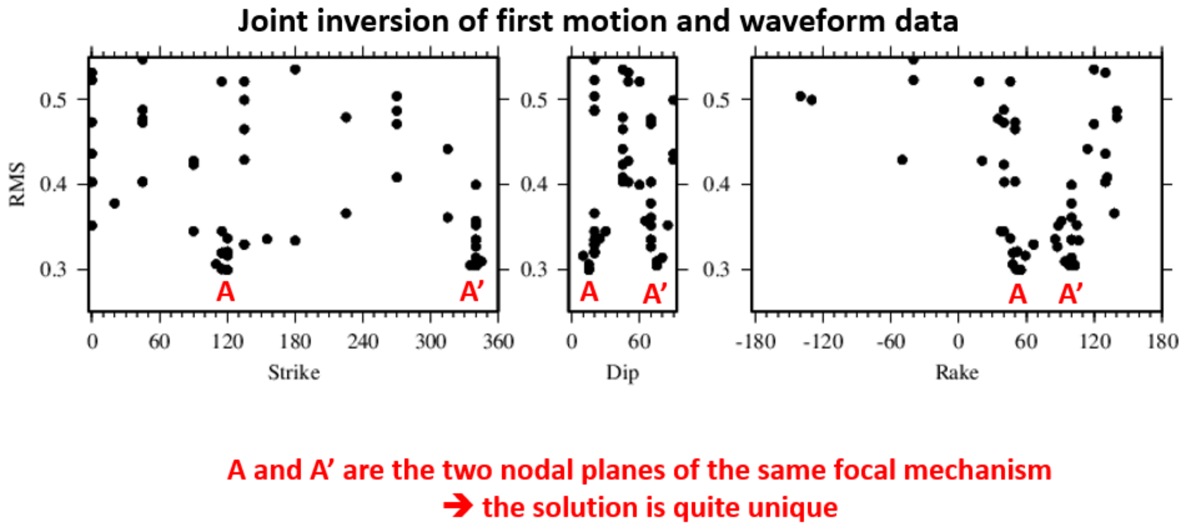

Figure S9. *Best solutions found at the end of the various grid search and the simulated annealing combinations for the joint inversion of first motion and waveform data, for the 2022/08/14 08:05:06 UTC Mw 2.5 event. Only the waveforms of the four nearest stations are used.*

Figure S7, S8, S9 are now included in the electronic supplement of the article, with the accompanying text.

- because I can tell from Fig 9 that the strike of this nodal plane is not consistent with local block movements.

In Figure 9 (top map) of the manuscript, we can observe that the main active thrust structures are oriented NW – SE, though their trace at the surface may show a varying orientation, as for instance just west of Tunari (the TCF) . The sinuous trace of the fault thrust front at the surface and the cross-section (lower part of Figure 9) strongly suggest the existence of a sub-horizontal, i.e. nearly flat, decollement, with an overall motion of the hanging wall to the NE. Since the strike of a horizontal plane is undefined, the strike of a sub-horizontal plane is poorly defined. For that reason, we believe that the strike values of the near-horizontal nodal plane of the two focal mechanisms analyzed in this paper, which we consider to correspond to the decollement, need not correspond to the NW-SE orientation of the relief. On the other hand, the second nodal plane, which is nearly vertical, i.e. with a strike well defined, *prescribes the slip direction of the hanging wall over the decollement, which is indeed in the expected NE direction as shown by Figures 7e and 8e of the manuscript. In that sense, we do not see incompatibilities between the strike, dip, rake parameters found and the expected bloc movements in the area.*

- In your location, do you consider the station elevation? From Fig 9, I think the elevation is not ignorable.

True, the range of elevation of the different stations varies from 376 m (SODC8) to 4264 m (SODC6), see table below. The elevation is taken into account in REAL for our refined location procedure. The first layer velocity parameters are applied to the positive depths.

Stat. Name	Longitud (°)	Latitud (°)	Elevation (m)
SODC0	-66.5354	-17.2344	3076
SODC1	-66.1085	-17.2426	3940
SODC2	-65.7857	-17.4667	2980
SODC3	-66.2215	-17.5046	3103
SODC4	-66.0738	-17.3715	2829
SODC5	-66.2124	-17.3028	3311
SODC6	-66.4288	-17.1668	4264
SODC7	-65.8727	-17.2058	2609
SODC8	-65.4711	-17.0300	376
SODC9	-65.5106	-17.6361	3340
SODC10	-66.3616	-17.6067	2867

- The title in the main text and supplemental material are different

We fixed it, and reported the title of the manuscript «Unveiling midcrustal seismic activity at the front of the bolivian altiplano, Cochabamba region »in the supplemental data material.

Reviewer C Comments (Round 3)

For author and editor

The authors have addressed all comments I have, and the current version is suitable for publication